Validation of aggregated risks models

Dacorogna, Michel ; Elbahtouri, Laila ; Kratz, Marie

Abstract: Validation of risk models is required by regulators and demanded by management and shareholders. Those models rely in practice heavily on Monte Carlo (MC) simulations. Given their complexity, the convergence of the MC algorithm is difficult to prove mathematically. To circumvent this problem and nevertheless explore the conditions of convergence, we suggest an analytical approach. Considering standard models, we compute, via mixing techniques, closed form formulas for risk measures as VaR or TVaR on a portfolio of risks, and consequently for the associated diversification benefit. The numerical convergence of MC simulations of those various quantities is then tested against their analytical evaluations. The speed of convergence appears to depend on the fatness of the tail of the marginal distributions; the higher the tail index, the faster the convergence. We also explore the behavior of the diversification benefit with various dependence structures and marginals (heavy and light tails). As expected, it varies heavily with the type of dependence between aggregated risks. The diversification benefit is also studied as a function of the risk measure, VaR or TVaR.

DOI: https://doi.org/10.1017/S1748499517000227

Annals of Actuarial Science, Vol. 12, part 2, pp. 433–454. © Institute and Faculty of Actuaries 2017
doi:10.1017/S1748499517000227
First published online 4 December 2017

Validation of aggregated risks models

Michel Dacorogna
DEAR-Consulting, Scheuchzerstrasse 160, 8057 Zurich, Switzerland

Laila Elbahtouri
SCOR, 5 Avenue Kléber, 75795 Paris, France

Marie Kratz*
*ESSEC Business School, CREAR Risk Research Center, Avenue Bernard Hirsch, BP50105, 95021 Cergy-Pontoise, France**

Abstract
Validation of risk models is required by regulators and demanded by management and shareholders. Those models rely in practice heavily on Monte Carlo (MC) simulations. Given their complexity, the convergence of the MC algorithm is difficult to prove mathematically. To circumvent this problem and nevertheless explore the conditions of convergence, we suggest an analytical approach. Considering standard models, we compute, via mixing techniques, closed form formulas for risk measures as Value-at-Risk (VaR) VaR or Tail Value-at-Risk (TVaR) TVaR on a portfolio of risks, and consequently for the associated diversification benefit. The numerical convergence of MC simulations of those various quantities is then tested against their analytical evaluations. The speed of convergence appears to depend on the fatness of the tail of the marginal distributions; the higher the tail index, the faster the convergence. We also explore the behaviour of the diversification benefit with various dependence structures and marginals (heavy and light tails). As expected, it varies heavily with the type of dependence between aggregated risks. The diversification benefit is also studied as a function of the risk measure, VaR or TVaR.

Keywords
Diversification (benefit); Heavy tail; Mixing technique; Model validation; Monte-Carlo simulation

1. Introduction

Risk and capital considerations are becoming central to the risk management of (re)insurance. In particular, the advent of risk-based solvency regulation like Solvency 2 in the European Union or the Swiss Solvency Test is bringing along many new requirements that push the industry in this direction. For instance, Solvency 2 Directive requires from (re)insurance companies to assess the capital needed to ensure their solvency. It is defined as the Value-at-Risk (VaR) of the own funds (the aggregated asset and liability risks) at a 1-year horizon subject to a confidence level of 99.5% over a 1-year period. The validation of this assessment is also required by the Directive and necessary for management to use the models for strategic decisions. It is not straightforward to validate results computed at such high threshold. Therefore, it should be tackled indirectly with various techniques.

*Correspondence to: ESSEC Business School, CREAR Risk Research Center, Avenue Bernard Hirsch, BP50105, 95021 Cergy-Pontoise, France. E-mail: kratz@essec.edu

This paper is a step in this direction, proposing ways to check both the convergence of Monte Carlo (MC) algorithm and the behaviour of diversification benefit.

Understanding the diversification benefit is essential to the business model of reinsurance companies. Diversification is at the heart of efficient risk management, capital optimisation and competitive pricing. Internal models are considered as the best solution to monitor the risk within the (re)insurer's portfolio and to compute the Solvency Capital Requirements, provided that the models are approved by the regulators. However, the authorities demand a validation of the whole modelling process and in particular of the techniques used. This implies thorough mathematical and best practice justifications of the choice of each distribution, assumption, parameter estimation, risk aggregation technique and so on (see Dacorogna, 2017). Most models are based on MC simulations of a large number of dependent risks. The convergence of such models is hard to prove mathematically and their stability, as a function of the number of simulations, is difficult to assess. The analytical approach is a way to measure the performance of the MC method, and potentially to replace it in some cases. Here, we propose this approach to improve our understanding and test the validity of model results.

Using standard models of the literature, we provide explicit expressions for the probability density function (pdf) of aggregated dependent risks, allowing to compute analytically risk measures as VaR or Tail Value-at-Risk (TVaR), and consequently the diversification benefit. This can be considered as a step towards the development of methods for model validation. To achieve this, we use mixing techniques (see Marshall & Olkin, 1988; Oakes, 1989) over risk parameter values. It is a standard tool in credibility theory, but less explored for actuarial dependent risks modelling. Whereas it has been used in ruin theory (see Albrecher *et al.*, 2011), we introduce it for risk measures and diversification benefit. We could have used an alternative technique that has been proposed in Constantinescu *et al.* (2011) and Hashorva & Ji (2014), where conditional models are seen as simple random scaling models. But our purpose here is to concentrate on the numerical analysis in order to test the performance of the MC method.

On the same line of research, we use explicit formulas for the diversification benefit to explore its behaviour as a function of the aggregation factor and the risk measures through models examples exhibiting dependence or independence, and light or heavy tailed marginals.

The paper is organised as follows. We describe in section 2 the framework of constructing Archimedean copulas using the mixing techniques, deriving analytical expressions for risk measures and diversification benefit for two combinations of dependence structure and marginal distributions (Pareto–Clayton and Weibull–Gumbel). Note that special efforts have been put in presenting in a consistent and clear way the mixing techniques approach. In section 3, the numerical convergence of MC simulations of risk measures and diversification benefit is tested against the analytical results for those two dependent models, Pareto–Clayton and Weibull–Gumbel. In section 4, we compare the behaviour of the diversification benefit of models, from independence to various dependence forms. We study it as a function of risk measures and aggregation factors. General conclusions follow in the last section 5.

2. Analytical Results Via the Mixing Techniques

A general method for constructing multivariate Archimedean copulas has been introduced by Oakes (1989) in the bivariate case and extended in the multivariate case by Marshall & Olkin (1988).

The idea behind this method is to use the mixing technique over a latent variable as a tool for dependence modelling. Introducing a latent variable to transform dependent variables into conditionally independent ones, allows to express the dependence between the variables as an Archimedean survival copula with parameter the latent variable, and to obtain the marginal distributions depending on this parameter. Namely, we have

The Oakes–Marshall–Olkin Theorem[1]

Let Θ be a positive random variable (rv) with cumulative distribution function (cdf) F_Θ and X_k, $k \geq 1$, be rvs such that

$$P(X_1 > x_1, ..., X_n > x_n \mid \Theta = \theta) = \prod_{k=1}^{n} H(x_k)^\theta \tag{1}$$

H being a positive function. The dependence model specified by (1) is a variant of the structure dependence generated by an Archimedean survival copula with generator $\phi = L_\Theta^{-1}$, where L_Θ denotes the Laplace transform of F_Θ:

$$P[X_1 > x_1, \ \dots \ , X_n > x_n] = L_\Theta\left(\sum_{i=1}^{n} L_\Theta^{-1}\left(\overline{F}_i(x_i)\right)\right) \tag{2}$$

where the marginal distributions F_i of X_i, $i = 1, \dots, n$, are defined by

$$\overline{F}_i(x) : = 1 - F_i(x) = L_\Theta(-\ln H(x)) \tag{3}$$

By means of this mixing technique, we can provide an explicit formula for the pdf of aggregated risks X_i, $i = 1, \dots, n$, of a dependence model. Finding the appropriate choice of the function H and the mixing parameter Θ to fit the marginals and the dependence model is the key step to derive via the Oakes–Marshall–Olkin theorem an explicit formula for the pdf.

With the explicit pdf f_{S_n} of the aggregate risk $S_n : = \sum_{i=1}^{n} X_i$, denoted by f_n when no possible confusion, we can derive the formulas for the risk measures and the diversification benefit. Recall that the diversification performance of a portfolio S_n is measured on the gain of capital when considering a portfolio instead of a sum of standalone risks. The capital is defined by the deviation to the expectation, and the diversification benefit (Bürgi *et al.*, 2008) at a threshold κ $(0 < \kappa < 1)$, by

$$D_\kappa(S_n) = 1 - \frac{\rho_\kappa(S_n) - \mathbb{E}(S_n)}{\sum_{i=1}^{n} (\rho_\kappa(X_i) - \mathbb{E}(X_i))} = 1 - \frac{\rho_\kappa(S_n) - \mathbb{E}(S_n)}{\sum_{i=1}^{n} \rho_\kappa(X_i) - \mathbb{E}(S_n)} \tag{4}$$

where ρ_κ denotes a risk measure at threshold κ. This indicator helps determining the optimal portfolio of the company since diversification reduces the risk and thus enhances the performance. By making sure that the diversification benefit is maximal, the company obtains the best performance for the lowest risk. However, it is important to note that $D_\kappa(S_n)$ is not a universal measure and depends on the number of risks undertaken and the chosen risk measure.

We will consider two examples of dependent models, presented in Marshall & Olkin (1988) (and, since, considered in various papers, as, e.g., in Albrecher *et al.*, 2011), that are useful in the

[1] Note that we choose to denote this result as *the Oakes–Marshall–Olkin Theorem*. It has been introduced and developed by Oakes for the bivariate case (see Oakes, 1989) and generalised for any n by Marshall & Olkin (1988).

reinsurance context. The first model is with Pareto marginals and a Clayton structure of dependence, which is standard in reinsurance context as it captures the dependence in the tail. The second one considers Weibull marginals and a Gumbel copula; it is an interesting alternative since it combines tail dependence with thin tail distributions. Note that further examples have been considered in a recent preprint (Sarabia *et al.*, 2017).

Throughout the paper, we will assume the same threshold κ for any quantity defined w.r.t. this threshold, hence we will omit it in the notation of those quantities.

2.1. Pareto marginals with Clayton survival copula

In this example, we consider a dependent model $X=(X_1,\dots,X_n)$ with marginals F_i $(i=1,\dots,n)$ (α,β)-Pareto (also called Pareto Lomax) distributed with $\alpha>1$, $\beta>0$, i.e. such that

$$\overline{F_i}(x) := 1-F_i(x) = \left(1+\frac{x}{\beta}\right)^{-\alpha}, \quad \forall x>0, \quad i=1,\dots,n \tag{5}$$

If $\beta=1$, we simplify the notation writing α-Pareto.

Recall that the quantile q_1 of order κ of a (α,β)-Pareto cdf is given by

$$q_1 = \mathrm{VaR}(\kappa) = \beta\left((1-\kappa)^{-1/\alpha}-1\right) \tag{6}$$

The dependence structure of the model is chosen as a survival copula C_θ of Clayton form with parameter $\theta>0$, defined on $[0,1]^n$ by

$$C_\theta(u_1,\cdots,u_n) = \varphi_\theta^{-1}\left(\sum_{i=1}^n \varphi_\theta(u_i)\right) = \left(\sum_{i=1}^n u_i^{-\theta}-(n-1)\right)^{-1/\theta}$$

with its generator φ (invariant to multiplication of the argument by a positive constant) given by

$$\varphi_\theta(t) = t^{-\theta}-1, \quad t\in[0\,;\,1] \tag{7}$$

First we need to compute the pdf of the aggregate risk $S_n=\sum_{i=1}^n X_i$ to evaluate the risk measures VaR and TVaR of S_n, then deduce the associated diversification benefit. Computing directly the pdf of S_n associated to our dependent model may be a difficult task, hence the choice of using the mixing technique to work with conditional independence. Assuming $\theta=1/\alpha$ to ease the computation, we obtain the following result.

Proposition 2.1 Consider the dependent model $X=(X_1,\dots,X_n)$ with marginals F_i $(i=1,\dots,n)$ (α,β)-Pareto distributed (defined in (5)), $\alpha>1$, $\beta>0$, and Clayton survival copula defined in (7) with parameter $\theta=1/\alpha>0$. Then the pdf f_n of the aggregate risk $S_n=\sum_{i=1}^n X_i$ is compound Gamma (or β of the second kind with parameters (n,α)) given, for $s>0$, by

$$f_n(s) = \frac{\beta^\alpha}{B(\alpha,n)}\times\frac{s^{n-1}}{(\beta+s)^{\alpha+n}} \tag{8}$$

B denoting the β function.

The choice $\theta=1/\alpha$ is a constraint of this model. However, it can be generalised to separate the dependence from the tail index. This is the subject of a forthcoming paper.

Having an explicit formula (8) for the pdf f_n of S_n, we can deduce its cdf F_{S_n} integrating f_n, and any risk measure based on F_{S_n}.

The value-at-risk of n at threshold κ, denoted $q_{\kappa,\,n}$ or q_n, is obtained via:

$$q_n = VaR_\kappa(S_n) = F^{\leftarrow}_{S_n}(\kappa) \tag{9}$$

$F^{\leftarrow}_{S_n}$ denoting the inverse of F_{S_n}. Note that using the relation between a β distribution of second kind and a β distribution, q_n may also be expressed in terms of the quantile of the β distribution with the same parameters, $B(n,\,\alpha)$,

$$q_n = \frac{VaR_\kappa(B(n,\,\alpha))}{1-VaR_\kappa(B(n,\,\alpha))}$$

The TVaR of S_n at threshold κ, defined by

$$TVaR_n = TVaR_\kappa(S_n) = \mathbb{E}[S_n \mid S_n \geq q_n]$$

(F_{S_n} being continuous), can be computed explicitly in terms of q_n, as done in the proposition below.

For completeness, let us compute the TVaR at threshold κ for a $(\alpha,\,\beta)$-Pareto cdf, assuming $\alpha > 1$ (for the existence of the TVaR) and $\beta > 0$. We obtain

$$\begin{aligned} TVaR_1 \quad &:= TVaR_\kappa(X) = \tfrac{\beta^\alpha}{(1-\kappa)\,(\alpha-1)} \times \tfrac{\alpha q_1+\beta}{(q_1+\beta)^\alpha}, \quad \text{where } q_1 \text{ is given in (6)}, \\ &i.e. \quad TVaR_1 = \beta\left(\tfrac{1}{(1-1/\alpha)\,(1-\kappa)^{1/\alpha}} - 1\right) \end{aligned} \tag{10}$$

Theorem 2.1 Considering the dependent Pareto–Clayton model X given in Proposition 2.1, the TVaR of the aggregate risk $S_n = \sum_{i=1}^n X_i$ at a threshold κ, $0 < \kappa < 1$, is given by

$$TVaR_n = \frac{\beta}{(1-\kappa)B(\alpha,n)} B\left(\left(1+\frac{q_n}{\beta}\right)^{-1};\, \alpha-1,\, n+1\right) \tag{11}$$

where $B(x;\,a,\,b)$ denotes the incomplete β function defined by

$$B(x\,;\,a,b) = \int_0^x u^{a-1}(1-u)^{b-1}du$$

If the shape parameter α of the Pareto margin is such that $\alpha \in \mathbb{N} \setminus \{0,\,1\}$, then $TVaR_n$ simplifies to

$$\begin{aligned} TVaR_n \quad &= \tfrac{n\beta}{(1-\kappa)(\alpha-1)}\left(1-\sum_{j=0}^{\alpha-2}\binom{n+\alpha-1}{j}p_n^j(1-p_n)^{n+\alpha-1-j}\right), \quad \text{where } p_n := \left(1+\tfrac{q_n}{\beta}\right)^{-1}, \\ &i.e. \quad TVaR_n = \tfrac{n\beta}{(1-\kappa)(\alpha-1)}\,\mathbb{P}[Y > \alpha-2] \end{aligned} \tag{12}$$

where Y follows a Binomial distribution $\mathcal{B}(n+\alpha-1,\,p_n)$.

In the bivariate case $n = 2$, (11) can be simplified as:

$$TVaR_2 = \frac{\beta^\alpha}{(1-\kappa)\,(\alpha-1)} \times \frac{\alpha(1+\alpha)\,q_2^2 + 2\beta\,(1+\alpha)\,q_2 + 2\beta^2}{(q_2+\beta)^{1+\alpha}} \tag{13}$$

Note that formulas (11) and (12) also hold for $n = 1$, resulting in (10).

Now we can deduce an explicit formula of the diversification benefit associated to our model, defined in (4), and denoted by D_n when choosing TVaR as risk measure, and D_n^* for VaR.

Corollary 2.1 Consider the dependent Pareto–Clayton model $\underset{n}{X} = (X_i,\ i \geq 1)$ given in Proposition 2.1. Then the diversification benefit of the aggregate risk $S_n = \sum_{i=1}^{n} X_i$ at a threshold κ, $0 < \kappa < 1$, and associated to the risk measure ρ, can be expressed as:

(i) For $\rho = \text{VaR}$:

$$D_n^* = 1 - \frac{(q_n - n\,\mathbb{E}(X))}{n\,(q_1 - \mathbb{E}(X))} = 1 - \frac{\frac{1}{n\beta} q_n - \frac{1}{\alpha - 1}}{(1-\kappa)^{-1/\alpha} - \frac{\alpha}{\alpha - 1}} \tag{14}$$

q_1 being defined in (6) and q_n in (9).

(ii) For $\rho = \text{TVaR}$:

$$D_n = 1 - \frac{\frac{(\alpha-1)}{n(1-\kappa)\,B(\alpha,n)}\, B\left(\left(1 + \frac{q_n}{\beta}\right)^{-1};\ \alpha - 1,\ n+1\right) - 1}{\alpha\left((1-\kappa)^{-1/\alpha} - 1\right)} \tag{15}$$

which simplifies, for $n = 2$, to

$$D_2 = 1 - \frac{\frac{\beta^{\alpha-1}}{2(1-\kappa)} \times \frac{\alpha(1+\alpha)\, q_2^2 + 2\beta(1+\alpha)\, q_2 + 2\beta^2}{(q_2+\beta)^{1+\alpha}} - 1}{\alpha\left((1-\kappa)^{-1/\alpha} - 1\right)} \tag{16}$$

2.2. Weibull marginals with Gumbel survival copula

We focus now our attention to another type of distribution associated with a different copula: Weibull marginals with Gumbel dependence. Gumbel dependence is also an Archimedean survival copula presenting asymmetric dependence with strong tail dependence although less asymmetric than the Clayton survival copula. The Weibull distribution is also used in insurance particularly for survival analysis or large claim occurrences, but the tail is usually less heavy than with certain Pareto distributions.

Consider a dependent model $X = (X_1, \ldots, X_n)$ with marginals F_i $(i = 1, \ldots, n)$ (c, τ)-Weibull distributed (with $c > 0$, $\tau > 0$), i.e. such that, for all $x \geq 0$,

$$\overline{F_i}(x) :\ = 1 - F_i(x) = e^{-cx^{\tau}} \tag{17}$$

When $c = 1$, it is called τ-Weibull.

Recall that the quantile q_1 of order κ $(0 < \kappa < 1)$ of a (c, τ)-Weibull cdf is given by

$$q_1 = \text{VaR}(\kappa) = \left(-\frac{\ln(1-\kappa)}{c}\right)^{1/\tau} \tag{18}$$

The dependence structure of the model is chosen as the survival copula of Gumbel form, with parameter $\theta \geq 1$, and generator φ defined by

$$\varphi_\theta(t) = (-\ln(t))^{\theta} \tag{19}$$

For simplicity of computations, we fix most parameters of the models, namely $\tau = 1/2$ and $\theta = 1/\tau$. Nevertheless, extension to $\theta = 1/\tau$ (but still with $\theta = 1/\tau$) has been recently developed in Sarabia *et al.* (2017).

Proposition 2.2 Consider the dependent model $X=(X_1,\dots,X_n)$ with marginals F_i $(i=1,\dots,n)$ (c,τ)-Weibull distributed with $c>0$ and $\tau=1/2$, and Gumbel survival copula with parameter $\theta=1/\tau$. Then the pdf f_n of the aggregate risk $S_n=\sum_{i=1}^{n}X_i$ at a threshold κ, $0<\kappa<1$, $n\geq 1$, is given, for $s>0$, by

$$f_n(s)=\frac{c}{2\sqrt{\pi}}\frac{\Gamma\left(n-\frac{1}{2}\right)}{\Gamma(n)}\,s^{-\frac{1}{2}}e^{-c\sqrt{s}}\,{}_1F_1\left(1-n,\,2-2n;\,2c\sqrt{s}\right) \tag{20}$$

where ${}_1F_1(a,\,b;\,x)$ is the Kummer confluent hypergeometric function defined on $\mathbb{R}$, with real parameters a, b, by ${}_1F_1(a,\,b;\,z)=\sum_{k=0}^{\infty}\frac{(a)_k}{(b)_k}\frac{z^k}{k!}$ where $(a)_k=\frac{\Gamma(a+k)}{\Gamma(a)}$ (see, e.g., Gradshteyn, 1988: 958 and Koll and Mohl 2011).

In particular, for $n=2$, (20) simplifies to

$$f_2(s)=\frac{c}{4}\left(\frac{1}{\sqrt{s}}+c\right)e^{-c\sqrt{s}},\quad s>0 \tag{21}$$

The pdf of S_n being now explicit, we can compute the expressions of the TVaR and the diversification benefit associated to our model for n risks.

Theorem 2.2 Consider the dependent Weibull–Gumbel model $X=(X_i,\,i\geq 1)$ given in Proposition 2.2. Then the TVaR of the aggregate risk $S_n=\sum_{i=1}^{n}X_i$ at a threshold κ, $0<\kappa<1$, can be expressed, for $n\geq 1$, as:

$$TVaR_n=\frac{2n\,e^{-c\sqrt{q_n}}}{(1-\kappa)c^2}\left(1+c\,q_n^{1/2}+\frac{c^2q_n}{2}+\frac{c^3\,q_n^{3/2}}{4\sqrt{\pi}}\sum_{k=2}^{n}\frac{\Gamma\left(k-\frac{3}{2}\right)}{k!}\,{}_1F_1(2-k,4-2k;\,2c\sqrt{q_n})\right) \tag{22}$$

with $q_n=F_{S_n}^{\leftarrow}(\kappa)$, F_{S_n} being defined in (20).

In particular we have, for $n=1$,

$$TVaR_1=\frac{2}{c^2}\left(1-\ln(1-\kappa)+\frac{1}{2}(\ln(1-\kappa))^2\right) \tag{23}$$

and, for $n=2$,

$$TVaR_2=\frac{4e^{-c\sqrt{q_2}}}{(1-\kappa)c^2}\left(1+c\,q_2^{1/2}+\frac{c^2q_2}{2}+\frac{c^3q_2^{3/2}}{8}\right) \tag{24}$$

Corollary 2.2 Consider the dependent Weibull–Gumbel model $X=(X_i,\,i\geq 1)$ given in Proposition 2.2. Then the diversification benefit of the aggregate risk $S_n=\sum_{i=1}^{n}X_i$ at a threshold κ $(0<\kappa<1)$, associated with the risk measure ρ, can be expressed as:

(i) For $\rho=\text{VaR}$, $D_n^*=1-\frac{q_u-n\mathbb{E}(X)}{n\,q_1-n\mathbb{E}(X)}=1-\frac{c^2(q_u-2nc)}{n\left((\ln(1-\kappa))^2-2c^3\right)}$

(ii) For $\rho=\text{TVaR}$,

$$D_n=1-\frac{e^{-c\sqrt{q_n}}\left(1+cq_n^{1/2}+\frac{c^2}{2}\,q_n+\frac{c^3}{4\sqrt{\pi}}\,q_n^{3/2}\sum_{k=2}^{n}\frac{\Gamma\left(k-\frac{3}{2}\right)}{k!}\,{}_1F_1\left(2-k,\,4-2k;\,2c\sqrt{q_n}\right)\right)-c^3(1-\kappa)}{(1-\kappa)\left(1-c^3-\ln(1-\kappa)+\frac{1}{2}(\ln(1-\kappa))^2\right)} \tag{25}$$

which simplifies, for $n=2$, to:

$$D_2 = 1 - \frac{e^{-c\sqrt{q_2}}\left(1 + cq_2^{1/2} + \frac{c^2}{2}\,q_2 + \frac{c^3}{8}\,q_2^{3/2}\right) - c^3(1-\kappa)}{(1-\kappa)\left(1 - c^3 - \ln(1-\kappa) + \frac{1}{2}(\ln(1-\kappa))^2\right)} \tag{26}$$

3. Testing the Quality of the Estimations Obtained Via MC simulations

The main benefit of explicit formulas is to provide exact answers for the risk measures and the diversification benefit, which are often estimated through MC simulations, which convergence is not well known. With explicit formulas, it is thus easy to check the quality of the estimations using MC. We do it on the two previous examples and proceed as follows.

On one hand, using the analytical expressions obtained in the previous section, we compute the TVaR and the diversification benefit for various values of the aggregation factor n, namely 2, 10 and 100 as illustrations. Those numbers will give us a benchmark for the MC simulations.

On the other hand, we run ten sets of simulations (changing the seed of the random generator) for each of the model parameters varying the number of simulations per run from 10,000 to 10 million. We report the average value over the ten sets of simulations and check that the standard deviation of those sets decreases, as expected, with the number of simulations.

Finally we compare those average values with the benchmark obtained with analytical formulas.

3.1. Pareto marginals with Clayton survival copula

We compute the TVaR and the diversification benefit (via Theorem 2.1 and Corollary 2.1) for various values of the aggregation factor n (2, 10, 100) and for a limited set of parameters. For the (α, β)-Pareto marginals, we fix $\beta = 1$.

Using the analytical expression (11) (and (12) when α is an integer $\neq 0, 1$) as benchmark, we check the convergence of the simulated TVaR, varying the parameter α of the tail index and the aggregation factor n. For the parameters, we are limited here to one free (α), as the Clayton survival copula parameter θ relates directly to the tail index α with $\alpha = 1/\theta$.

First, we explore a case with very heavy tail, $\alpha = 1.1$, which corresponds to tails seen for earthquake distributions, and a relatively strong dependence ($\theta \approx 0.91$). Then we look at the case $\alpha = 2$ with moderate heavy tail, followed by that of a moderate tail $\alpha = 3$, which also means here a moderate dependence $\theta = 1/3$. We run ten sets of simulations (changing the seed of the random generator) for each of these parameters varying the number of simulations per run from 10,000 to 10 millions. We report here the average value over the ten sets of simulations. It is worth noticing that the standard deviation of those sets decreases, as expected, with the number of simulations. For the TVaR computed for $n = 2$, the standard deviation of the ten sets varies from 32% to 4% for $\alpha = 3$ and from 57,968% to 247% for $\alpha = 1.1$ going from 100,000 to 10 millions simulations (except for $n = 100$, where we stop at one million due to computer limitations). As expected, the convergence is much slower in the case of a fatter tail. Moreover, for extremely fat tail (α close to one), it does not converge even for ten millions simulations. A similar behaviour can be observed for $n = 10$. Beside the gain in precision, the analytical formula can be numerically evaluated 40 times faster,

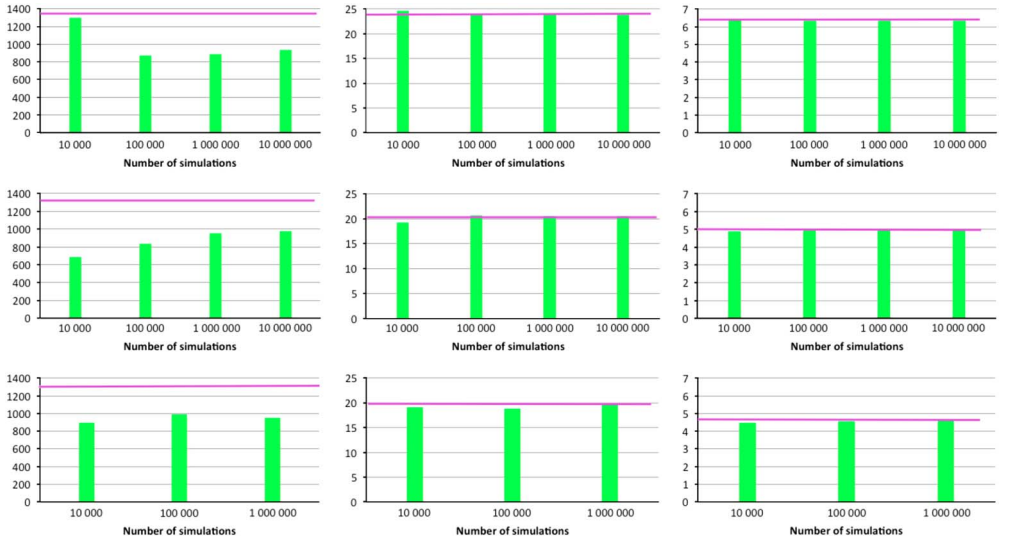

Figure 1. Convergence of the TVaR of S_n at 99.5% for $\alpha = 1.1$, 2, 3 from left to right, for an aggregation factor $n = 2$, 10, 100 from up to down. The purple line corresponds to the analytical value and the green plots are the average values obtained from the Monte Carlo (MC) simulations. The y-scale gives the normalised TVaR ($TVaR_n/n$) and is the same for each column.

respectively, 580 times faster (for $\alpha = 2$ and $n = 10$, respectively, $n = 100$ for one million simulation) than the estimation given by MC simulations.

For comparability reasons, we present in Figure 1 the normalised TVaR, $TVaR_n/n$, for the various n's. On the figure, we see that:

- The normalised TVAR of S_n, $TVaR_n/n$, decreases as n increases
- The TVaR decreases as α increases
- The rate of convergence of $TVaR_n/n$ increases with n
- The heavier the tail, the slower the convergence
- In the case of very heavy tail and strong dependence ($\alpha = 1.1$ and $\theta = 0.91$), we do not see any satisfactory convergence, even with ten million simulations, and for any n. At ten million simulations, the value of TVaR is still 24% lower than the theoretical value.
- When $\alpha = 2$, 3, the convergence is good from one million, 100,000 simulations onwards, respectively.

In Table 1, we note that the convergence of the TVaR of S_n for the Pareto parameter $\alpha = 2$ is good, with a relative error going from 0.38% to 0.05% when n goes from 2 to 100, and for one million simulations. Similar relative errors are obtained for $\alpha = 3$. At ten million simulations, for $\alpha = 1.1$ and $n = 10$, the TVaR is still underestimated by 25%, which is unacceptable for an evaluation of the solvency capital. It seems clear that MC will not give satisfactory answers with reasonable number of simulations. The way out is to resort to explicit formula as derived in this paper to obtain credible values for the capital. However, increasing the number of aggregation improves the convergence. Measuring the changes with 1 million simulations, we see that from $n = 2$, we have an underestimation of 50% that decreases to 27% for $n = 10$ and $n = 100$.

Table 1. Relative errors (when comparing results obtained by Monte Carlo and analytical ones) of the $TVaR_n$ and the diversification benefit D_n for S_n, at 99.5% and for various α, as a function of the aggregation factor n computed with one million simulations.

	$n=2$ (%)	$n=10$ (%)	$n=100$ (%)
$\alpha=1.1$			
$TVaR_n$	−33.3	−27.3	−26.9
D_n	1,786	742	653
$\alpha=2$			
$TVaR_n$	0.38	0.14	0.05
D_n	−2.61	−0.44	−0.14
$\alpha=3$			
$TVaR_n$	0.30	0.14	−0.10
D_n	−1.30	−0.25	0.15

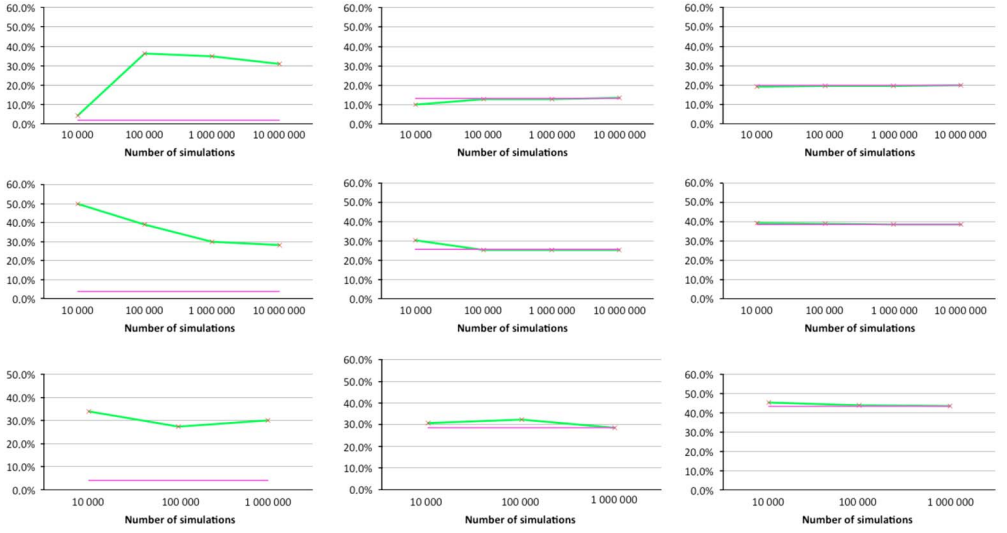

Figure 2. Convergence of the diversification benefit of S_n (associated with TVaR at 99.5%) for $\alpha=1.1$, 2, 3 (from left to right), for an aggregation factor $n=2$, 10, 100 (from up to down). The purple lines are for the analytical values and the green ones are the average values obtained from the Monte Carlo (MC) simulations. The y-scale is the same for all the graphs, for fair comparison.

Let us now analyse the diversification benefit $D_\kappa(S_n)$ associated with TVaR, denoted by D_n. We use the results (16) and (15) to compute it explicitly and then check the convergence of the MC. The same parameter set and the same simulations are used to produce the numbers displayed in Figure 2. As expected, the convergence of the diversification benefit follows a similar pattern as for the TVaR. Indeed, we see in Figure 2 that:

- The diversification benefit D_n of S_n increases with n
- D_n increases with α
- The rate of convergence of D_n increases with n
- The heavier the tail, the slower the convergence

- In the case of very heavy tail and strong dependence ($\alpha = 1.1$ and $\theta = 0.91$), we do not see any satisfactory convergence, even with ten million simulations, and for any n
- When $\alpha = 2$, 3, the convergence is good from one million, 100,000 simulations onwards, respectively.

For $\alpha = 1.1$, the convergence is poor. In this case, we see an overestimation of the diversification benefit that is above 600% for all aggregation factors including $n = 100$ (see Table 1). This significant overestimation may seem high, however, the impact is not material since we are talking here of very small diversification benefit of the order of a few percents (3.6% for $n = 10$, for instance). We see in Figure 2 that the diversification benefit does not converge in all cases. Obviously, ten millions simulations are not sufficient to ensure a good convergence. The reason is that MC samples the space evenly, whereas, in this case, we would need much more points in the tails to see a good convergence.

In Table 1, we present the results of a convergence study as a function of the aggregation factor. The number of simulations is fixed at one million and we vary the aggregation factor n. We see a decreasing estimation error by MC when increasing the aggregation factor, with small errors for $\alpha = 3$ and 2 and substantial errors for very fat tails and strong dependence. In the latter, we also see a systematic underestimation of the TVaR and an overestimation of the diversification benefit, whatever the aggregation factor. Such a fact is not surprising for very fat tails (α close to 1). With the thinner tails and lower dependence, MC has a tendency to overestimate the TVaR and underestimate the diversification benefit except for $n = 100$. Note that the error decrease is large between 2 and 10, but much smaller afterwards.

3.2. Weibull marginals with Gumbel survival copula

Here we use Theorem 2.2 and Corollary 2.2 to compute the TVaR and the diversification benefit, for various values of the aggregation factor n (2, 10, 100). Recall that in this example, $\theta = 2 = 1/\tau$, which means that the tail index is fixed to $\alpha = 2$. Only the scaling parameter c of the Weibull marginals might be modified, but for the sake of simplicity, we fix it to 1.

As in the previous example, we compare the values obtained analytically with those obtained via MC simulations, considering ten sets of simulations for our set of parameters, varying the number of simulations per run from 10,000 to ten millions (except for $n = 100$, where we stop at one million due to computer limitations). We report here the average values of the ten sets and verify that the standard deviation is decreasing with the number of simulations.

We present in Figure 3 the normalised TVaR, $TVaR_n/n$ (as we do it for Pareto–Clayton), for the three cases $n = 2$, 10, 100. We observe that:

- $TVaR_n/n$ decreases with n.
- $TVaR_n/n$ decreases faster between $n = 2$ and $n = 10$ (–29%) than between $n = 10$ and $n = 100$ (–12%).
- The rate of convergence is good in all the cases and the deviation reaches <1% already with 100,000 simulations.

On the figure, the convergence is very clear. The absolute value of the relative error is <2% for both risk measures and for all n's. The convergence, reached already with 100,000 simulations, is faster than the one obtained with the Pareto–Clayton model. This is explained by the fact that, for $0 < \tau < 1$, the Weibull marginals are moderately heavy tailed and the Gumbel survival copula has a

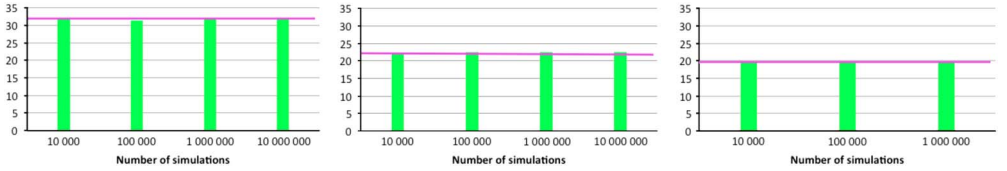

Figure 3. Convergence of the normalised TVaR of S_n, $TVaR_n/n$, at 99.5% for $c=1$, $\tau=\frac{1}{2}$ and $\theta=2$, for an aggregation factor $n=2$, 10, 100 from left to right. The purple line corresponds to the analytical value and the green plots are the average values obtained from the Monte Carlo (MC) simulations. The y-axis is the same for the three plots.

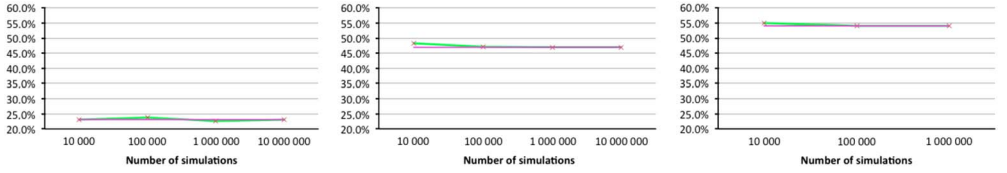

Figure 4. Convergence of the diversification benefit D_n of S_n (associated with TVaR at 99.5%) for $c=1$, $\tau=\frac{1}{2}$ and $\theta=2$, for an aggregation factor $n=2$, 10, 100 from left to right. The purple lines are for the analytical values and the green ones are the average values obtained from the Monte Carlo (MC) simulations. The y-axis is the same for the three plots, for fair comparison.

much weaker dependence in the tail compared to the Clayton survival copula; indeed recall that a Clayton copula has lower tail dependence (hence the flipped Clayton copula, as here, will exhibit only upper tail dependence), whereas the Gumbel one has upper tail dependence (the survival copula of Gumbel form will then exhibit only lower tail dependence). Thus, the simulation requires less points to model accurately the behaviour in the tails. Increasing the number of aggregation improves the convergence. Indeed, when measuring the variation with one million simulation for the TVaR, we see that for $n=2$ we have an underestimation of 0.65%, which decreases to 0.18% for $n=100$. Beside the gain in precision, the analytical formula can be numerically evaluated 65 times faster, respectively, 75 times faster (for $n=10$, respectively, $n=100$ for one million simulation), than the estimation given by MC method. Moreover, MC method cannot be used for higher number of aggregation (i.e. $n=10{,}000$) due to lack of system memory, while it is of course feasible for the analytical formula.

In Figure 4, we present the results for the diversification benefit D_n, when choosing TVaR as risk measure. Similar comments hold for D_n as for $TVaR_n/n$. The convergence is already very good for 10,000 simulations. We also observe, as expected, that D_n increases with n, as it is the case for the Pareto–Clayton model.

4. Diversification Benefit as a Function of the Aggregation Factor and the Risk Measures, for Various Types of Models

The diversification benefit is an important parameter for determining the efficiency of the use of capital of an insurance company. It plays a crucial role in the new risk disclosure of companies. Unfortunately, it is a quantity that has attracted little attention of researchers because it is not a universal measure: it depends on the number of underlying risks in the definition, as well as on the choice of the underlying risk measure (see Emmer *et al.*, 2015). However, this parameter can be

studied as a function of the number of risks, as well as of the risk measure. It is what we aim at doing here, addressing in particular the question of the impact of the existence of dependence among risks on the behaviour of the diversification benefit. In a previous paper (see Busse *et al.*, 2014), we explored analytically the limit of diversification when introducing a systematic risk in the model through a latent rv common to all risks in the portfolio. Here we tackle again the problem, but in a different way, using various dependence models and, as in Busse *et al.* (2014), comparing it to the independent case. We also study the effect of light and heavy tailed marginals.

When assuming dependence, we consider our two previous models, Pareto–Clayton and Weibull–Gumbel. We add two cases, one with independent Pareto rv's, to evaluate the impact of the dependence, and one considering a multivariate Gaussian model to evaluate the impact of the tail thickness.

All chosen examples share the existence of anaytical expressions for the risk measures and, as consequence, for the diversification benefit. We can then compare the behaviour of the diversification benefit as a function of the aggregation factor n (from $n=2$ to 10,000) and the risk measures, TVaR and VaR. Since both are used in practice, it is interesting to see if the choice of risk measure has an influence on the diversification benefit.

Note that for comparison purpose, we choose the Pareto–Clayton case with $\alpha=2$, since it corresponds to the tail index of the Weibull–Gumbel model. For completness, we also recall briefly the evaluation of the diversification benefit in the two additional examples, independent Pareto and mutivariate Gaussian risks.

4.1. Independent Pareto rv's case with asymptotic threshold ($\kappa \nearrow 1$)

To avoid computations, we are going to look only for approximations when considering extreme quantiles (i.e. when the threshold κ tends to 1), for which Feller's result (see Feller, 1966) is available. For sharper and not necessarily asymptotic results, computations could be done using the Normex method (see Kratz, 2014).

Feller has shown that the tail distribution of the sum of independent rv's with regularly varying (RV_α) tail distribution is asymptotically RV_α. Applying this result when considering iid Pareto-(α, β) rv's provides

$$q_n = VaR_\kappa(S_n) = F^{\leftarrow}_{S_n}(\kappa) \underset{\kappa\to 1}{\sim} \beta\left(\left(\frac{1-\kappa}{n}\right)^{-1/\alpha} - 1\right) \tag{27}$$

and

$$TVaR_\kappa(S_n) \underset{\kappa\to 1}{\sim} \frac{n\beta^\alpha}{(1-\kappa)(\alpha-1)} \times \frac{\alpha q_n + \beta}{(q_n+\beta)^\alpha} \underset{\kappa\to 1}{\sim} \beta\left(\frac{\alpha}{\alpha-1}\left(\frac{1-\kappa}{n}\right)^{-1/\alpha} - 1\right) \tag{28}$$

It gives back the well known asymptotic relation $TVaR_\kappa / VaR_\kappa \to \alpha/(\alpha-1)$ as $\kappa \to 1$.

Now let us look at the diversification benefit for high threshold κ. When choosing the VaR as risk measure (q_1 satisfying (6)), we have

$$D^*_n(\kappa) = 1 - \frac{q_n - \frac{n\beta}{\alpha-1}}{n\left(q_1 - \frac{\beta}{\alpha-1}\right)} \underset{\kappa\to 1}{\sim} 1 - \frac{n^{\frac{1}{\alpha}-1}(1-\kappa)^{-1/\alpha} - \frac{1}{n} - \frac{1}{\alpha-1}}{(1-\kappa)^{-1/\alpha} - 1 - \frac{1}{\alpha-1}} \tag{29}$$

and for the TVaR,

$$D_n(\kappa) \underset{\kappa\to 1}{\sim} 1-\frac{\frac{\alpha}{\alpha-1}\, n^{\frac{1}{\alpha}-1}\,(1-\kappa)^{-1/\alpha}-\frac{1}{n}-\frac{1}{\alpha-1}}{\frac{\alpha}{\alpha-1}(1-\kappa)^{-1/\alpha}-1-\frac{1}{\alpha-1}} \tag{30}$$

from which we deduce the asymptotic limit as $\kappa \to 1$, which is the same for both risk measures, as expected:

$$\lim_{\kappa\to 1} D_n^*(\kappa) = \lim_{\kappa\to 1} D_n(\kappa) = 1-n^{-(1-1/\alpha)} \tag{31}$$

which tends to 1 as $n \to \infty$.

Note that in the Gaussian case, D_n (or D_n^*), converges also to 1 as $n \to \infty$ with a rate of convergence of $n^{1/2}$, and not only for high threshold κ.

4.2. Multivariate Gaussian distribution

Consider the Gaussian vector $(X_i)_{i=1,\dots,n}$ with expectation vector $(\mu_i)_{i=1,\dots,n}$, $\mu_i=\mu$, and (non-negative definite) covariance matrix $\Gamma=(\gamma_{ij})_{1\le i,\, j\le n}$ such that $\gamma_{ii}=\sigma^2$, $\forall i$. Then $S_n=\sum_{i=1}^n X_i$ is normally distributed with mean $n\mu$ and variance $n\sigma^2+2\sum_{1\le i<j\le n}\gamma_{ij}$. Hence, with the notation $r_{ij}=\mathrm{corr}(X_i, X_j)$ and ϕ, Φ, for the standard normal pdf, cdf, respectively, we can write

$$q_n = VaR_\kappa(S_n) = n\mu+\Phi^{-1}(\kappa)\sqrt{n}\sigma\sqrt{1+\frac{2}{n}\sum_{1\le i<j\le n} r_{ij}} \le n\left(\mu+\Phi^{-1}(\kappa)\,\sigma\right) = n\,q_1$$

whereas, for the TVaR,

$$TVaR_\kappa(S_n) = n\mu+\frac{\phi\left(\Phi^{-1}(\kappa)\right)}{1-\kappa}\sqrt{n}\,\sigma\sqrt{1+\frac{2}{n}\sum_{1\le i<j\le n} r_{ij}} \le n\left(\mu+\frac{\phi\left(\Phi^{-1}(\kappa)\right)}{1-\kappa}\,\sigma\right) = n\,TVaR_\kappa(X)$$

We deduce that

$$D_n = D_n^* = 1-\frac{1}{\sqrt{n}}\sqrt{1+\frac{2}{n}\sum_{1\le i<j\le n} r_{ij}} \quad (\ge 0)$$

The diversification benefit can tend to any constant between 0 and 1, as $n \to \infty$, whenever there exists a linear dependence between the components. For instance, if $r_{ij}=r\neq 0$, $\forall i\neq j$, the diversification benefit reduces to

$$D_n = D_n^* = 1-\frac{1}{\sqrt{n}}\sqrt{1+(n-1)\,r}$$

when $r=1$ (full comonotonicity), $D_n=D_n^*=0$, whereas for $r=0$ (independent case), we obtain $D_n=D_n^*=1-n^{-1/2}$, which is also the limit (as $\kappa \to 1$) of the diversification benefit given in (31) for independent α-Pareto rv's with $\alpha=2$.

Let us compare the diversification benefit as a function of n, for both TVAR and VaR, for high threshold $\kappa=99.5\%$, considering the following cases: (i) independent α-Pareto with $\alpha=2$, (ii) α-Pareto margins and survival Clayton (θ) copula (Pareto–Clayton) with $\alpha=2=1/\theta$, (iii) Gaussian margins and Gaussian copula (Gaussian–Gaussian) with (linear) correlation $r=0.42$ estimated on the previous Pareto–Clayton model, (iv) τ-Weibull margins and Gumbel (θ) copula (Weibull–Gumbel) with $\theta=2=1/\tau$ and (v) Gaussian–Gaussian with correlation $r=0.39$ estimated on the Weibull–Gumbel model.

Table 2. Analytical diversification benefit of S_n, D_n and D_n^*, at 99.5% as a function of the risk measures TVaR and VaR, respectively, and of the aggregation factor n.

	$n=2$ (%)	$n=10$ (%)	$n=100$ (%)	$n=1{,}000$ (%)	$n=10{,}000$ (%)
Independent Pareto					
$D_n=D_n^*$	29.3	68.4	90.0	96.8	99.0
Pareto–Clayton					
D_n	13.2	25.5	28.6	29.0	29.0
D_n^*	12.9	25.2	28.3	28.6	28.7
D_n / D_n^*	1.021	1.014	1.012	1.012	1.012
Gaussian–Gaussian					
$r=0.42$ (Clayton)					
$D_n=D_n^*$	15.7	30.9	34.7	35.1	35.2
Weibull–Gumbel					
D_n	23.1	47.0	54.1	54.8	54.9
D_n^*	19.6	40.4	46.5	47.2	47.2
D_n / D_n^*	1.179	1.162	1.163	1.163	1.163
Gaussian–Gaussian					
$r=0.39$ (Gumbel)					
$D_n=D_n^*$	16.6	32.8	37.1	37.5	37.5

Note: Tail index $\alpha=2$, survival Clayton parameter $\theta=1/2$, Weibull $\tau=1/2$, Gumbel parameter $\theta=2$.

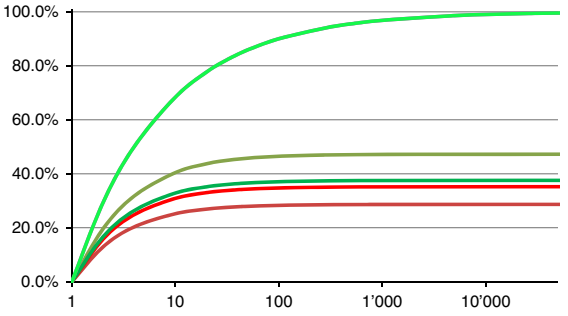

Figure 5. Plot of the diversification benefit D_n^* (y-axis) as a function of n (x-axis), for VaR_κ with $\kappa=99.5\%$. The curves from bottom to up correspond, respectively, to Pareto ($\alpha=2$)-Clayton ($\theta=1/2$), Gauss-Gauss (Clayton; $r=0.42$), Gauss-Gauss (Gumbel; $r=0.39$), Weibulll ($\tau=1/2$)-Gumbel ($\theta=2$), independent Pareto ($\alpha=2$) and Gauss models.

Note that the linear correlation coefficients r are estimated from the realisation of the simulated model. Moreover, the number of simulations is chosen such that the Kendall-τ estimate corresponds to the theoretical one. Recall that the theoretical value of τ, as a function of the θ parameter of the copula, is known for both Clayton, with $\tau=\frac{\theta}{\theta+2}$, and Gumbel, with $1-\frac{1}{\theta}$. The convergence is then reached with 100,000 simulations.

In Table 2, we present the results obtained on D_n and D_n^* using analytical formulas for all of them, and not simulations. Note that it completes the numerical application for the analytical diversification benefit done in the previous sections with the TVaR only. Moreover, the functions D_n^* of n are plotted in Figure 5 for the various models.

We show in Table 2 and Figure 5 that:

- The dependence has a significant impact on the evolution of the diversification benefit with n, as expected. In the case of dependence, the diversification benefit levels off rapidly, already at $n = 100$, while for the independence case, it still increases from $n = 100$ to $n = 10{,}000$.
- The evolution of the diversification benefit is similar for the two underlying structures, increasing and levelling off with n, for both. For example, in the dependent case, from $n = 2$ to 10, it increases approximately by a factor 2 (1.9 for Pareto–Clayton and 2.0 for Weibull–Gumbel), regardless of the risk measure used to compute the diversification benefit.
- Although both evolutions are the same from $n = 2$ to 100, the Pareto–Clayton model exhibits a stronger tail dependence and a fatter tail than the Weibull–Gumbel one. This appears through the difference of the diversification benefit between the two models, almost twice bigger for the second.
- We see in the table that the diversification benefit computed with the VaR is always lower than the one computed with the TVaR. This is to be expected, since the latter risk measure takes into account the whole distribution beyond a fixed threshold (for instance, 99.5% in this case), whereas the VaR only looks at one point of the distribution. However, the difference is more pronounced for the Weibull–Gumbel model than the Pareto–Clayton one, with a ratio of the order of 1.2 instead of 1.01, respectively.
- The ratio D_n / D_n^* stabilises from $n = 10$ onwards, with 1.012 for Pareto–Clayton and 1.163 for Weibull–Gumbel. Starting at $n = 100$, the diversification benefit does not change anymore. It is interesting to note that the more dependence in the tails (Clayton copula), the less difference between the diversification benefit measured with TVaR or VaR. Nevertheless, this difference is never very large (<20%).
- Comparing the results obtained with the Pareto–Clayton model and the corresponding Gaussian–Gaussian one, emphasises that the tail dependence associated with a fat tail limits the diversification benefit. The diversification benefit limit is 35.2% for the Gaussian case, while it is only 29% in the Pareto–Clayton case.
- The comparison leads to a different result in the Weibull–Gumbel case and the Gaussian one. Here we see much stronger diversification benefit for the Weibull–Gumbel model than for the Gaussian–Gaussian one. The intuition to explain this observation being less obvious than with the previous model, varying the parameters (as in Sarabia *et al.*, 2017) would help to analyse this effect. Let us just note that the linear correlation implied by $\theta = 2$ is quite high in this case.
- As expected, 100% diversification benefit is reached in the independent case only.

5. Conclusion

The purpose of this study is to explore new ways of validating internal models. As mentioned in the introduction, it is not possible to validate statistically a VaR at a threshold of 99.5%, which represents a probability of 1 over 200 years. Thus, we need to resort to indirect ways, testing various aspects of the model. Here we test the convergence of MC algorithms for aggregated risks in presence of both heavy tailed marginals and non-linear dependence in the tails. We also explore the behaviour of the diversification benefit as a function of: the number of risks, the type of marginal distributions, the dependence structure and the choice of risk measure used to evaluate it. This is done after having derived explicit formulas for both the risk measures and the diversification benefit.

To test the MC algorithm, we consider two standard examples of dependence structure and marginal distributions, Pareto–Clayton and Weibull–Gumbel. We obtain via mixing techniques explicit

formula for the aggregated pdf for both examples. By varying the number of simulations from 10,000 to 10 millions, we observe that for Pareto(α) – Clayton($1/\alpha$), a minimum of 100,000 simulations is needed to obtain a good convergence for moderately heavy tails and moderate tail dependence ($\alpha = 2$ or 3). For very heavy tails and moderate tail dependence (e.g. $\alpha = 1.1$), we do not reach convergence, even with ten millions simulations, showing that the results obtained by MC method for this kind of distributions must be considered carefully. The second case is more limited since the Gumbel copula parameter is fixed to $\theta = 2$. However, we can conclude that the convergence with a Gumbel copula is faster than with a Clayton survival copula, since 100,000 simulations are enough for moderate heavy tailed Weibull marginal distribution.

When studying the behaviour of the analytical diversification benefit of S_n as a function of the aggregation factor n, we observe that the dependence structure matters more for the diversification benefit than the type of marginals do. It is interesting to note that in the independent case, the diversification benefit is identical in the limit, when the threshold tends to 1, for Pareto with $\alpha = 2$ or Gaussian risks, despite the fact that Pareto has a heavier tail. We see that Gaussian marginals with Clayton or Gumbel copula have a lower diversification benefit than independent Pareto risks, or than Weibull marginals with Gumbel copula. With $n = 100$, the diversification benefit saturates in all the cases except for independent Gaussian (or Pareto $\alpha = 2$) risks; in this latter case, it gets close to 100% only when $n > 10{,}000$. Another interesting effect is that the diversification benefit depends little on the choice of risk measure for Pareto–Clayton (change around 2%), while in the Weibull–Gumbel case, the change can go up to 20%. It is a question we will explore further when considering α and θ independently.

Replacing MC simulations by explicit expressions is a considerable gain in precision and time. Indeed, the analytical formulas are estimated to be 40–500 times faster than MC method, with the time difference increasing with increasing number of aggregated variables. Moreover, with the analytical formula, we can estimate the TVaR even for very heavy tailed distributions whenever MC method fails even for ten millions simulations. Explicit formulas allow us to explore the aggregation behaviour of the risk measures and the diversification benefit. It is a precious tool for validating results of internal models, which are based on MC simulations. This study is a first step towards a general approach where we could choose the tail index and the copula parameter independently. It will be the object of further investigation.

Acknowledgement

Partial support from RARE-318984 (an FP7 Marie Curie IRSES Fellowship) is kindly acknowledged.

References

Albano, M., Amdeberhan, T., Beyerstedt, E. & Moll, V.H. (2011). The integrals in Gradshteyn and Ryzhik. Part 19: the error function. *SCIENTIA Series A: Mathematical Sciences*, **21**, 25–42.

Albrecher, H., Constantinescu, C. & Loisel, S. (2011). Explicit ruin formulas for models with dependence among risks. *Insurance: Mathematics and Economics*, **48**(2), 265–270.

Bürgi, R., Dacorogna, M.M. & Iles, R. (2008). Risk aggregation, dependence structure and diversification benefit. In D. Rösch & H. Scheule (Eds.), *Stress Testing for Financial Institutions* (pp. 265–306), Riskbooks, Incisive Media, London.

Busse, M., Dacorogna, M.M. & Kratz, M. (2014). The impact of systemic risk on the diversification benefits of a risk portfolio. *Risks*, **2**, 260–276.

Constantinescu, C., Hashorva, E. & Ji, L. (2011). Archimedean copulas in finite and infinite dimensions – with application to ruin problems. *Insurance: Mathematics and Economics*, **49**, 487–495.
Dacorogna, M. (2017). Approaches and techniques to validate internal model results. Available online at the address https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2983837 [accessed May 2017].
Dacorogna, M., Elbahtouri, L. & Kratz, M. (2015–2016). Explicit diversification benefit for dependent risks, ESSEC WP 1522 (2015) & SCOR Paper 38.
Emmer, S., Kratz, M. & Tasche, D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, **18**(2), 31–60.
Feller, W. (1966). *An Introduction to Probability Theory and its Applications* (Vol. II). Wiley, New-York.
Gradshteyn, I.S. (1988). Tables of Integrals, Series, and Products, Academic Press, San Diego.
Hashorva, E. & Ji, L. (2014). Random shifting and scaling of insurance risks. *Risks*, **2**, 277–288.
Kohl, K.T. & Moll, V.H. (2011). The integrals in Gradshteyn and Ryzhik. Part 20: hypergeometric functions. *SCIENTIA Series A: Mathematical Sciences*, **21**, 43–54.
Kratz, M. (2014). *Normex, a new method for evaluating the distribution of aggregated heavy tailed risks. Application to risk measures*. Extremes **17**(4), Special issue on *Extremes and Finance* (Guest Ed. P. Embrechts), 661–691.
Marshall, A.W. & Olkin, I. (1988). Families of Multivariate Distributions. *Journal of the American Statistical Association*, **83**, 834–841.
Oakes, D. (1989). Bivariate Survival Models Induced by Frailties. *Journal of the American Statistical Association*, **84**, 487–493.
Sarabia, J., Gómez-Déniz, E., Prietoa, F. & Jordá, V. (2017). Aggregation of Dependent Risks in Mixtures of Exponential Distributions and Extensions. Available online at the address https://arxiv.org/abs/1705.00289 [accessed May 2017].

Appendix

We provide the proofs of the analytical results given in section 2. For further details, we refer to Dacorogna *et al.* (2015–2016).

A. Proofs of the results given for the Pareto–Clayton model

- *Proof of Proposition 2.1*

The proof relies mainly on the application of the Oakes–Marshall–Olkin Theorem, which makes the computations of f_n easier.

As given in Marshall & Olkin (1988), choosing $H(x) = e^{-x}$ for $x > 0$, and the latent rv Θ a Gamma $\Gamma(\alpha; \beta)$ rv with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$ (or scale parameter $1/\beta$) with pdf $f_\Theta(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}_{(x \geq 0)}$, allows to check that the model defined by (2) and (3), which corresponds to our model, satisfies (1) by the Oakes–Marshall–Olkin Theorem (see Dacorogna *et al.*, 2015–2016).

Since conditioning X by Θ transforms the dependent risks into independent conditional ones, we write the pdf f_n of S_n as

$$f_n(s) = \int_0^\infty f_{S_n \mid \Theta}(s) f_\Theta(\theta) d\theta$$

With the choice of H and Θ, our model satisfies (1), hence the conditional rv's $X_i \mid \Theta = \theta$, $i = 1, \dots, n$, are i.i.d. exponentially distributed with parameter θ. We deduce that their conditional sum $S_n = \sum_{i=1}^{n} X_i \mid \Theta = \theta$ is $\Gamma(n, \theta)$-distributed. So we obtain immediately that the pdf of S_n is the one of a compound γ distribution and is given by

$$f_n(s) = \frac{s^{n-1}}{\Gamma(n)} \int_0^\infty \theta^n e^{-s\theta} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} d\theta = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\Gamma(n)} \times \frac{\beta^\alpha s^{n-1}}{(\beta+s)^{\alpha+n}}$$

hence the result (8).

- *Proof of Theorem 2.1*

Using the definition of TVaR and Proposition 2.1, we obtain

$$TVaR_n = \frac{1}{1-\kappa} \int_{q_n}^\infty s f_n(s)\, ds = \frac{\beta}{(1-\kappa)B(\alpha, n)} \int_{\frac{q_n}{\beta}}^\infty s^n (1+s)^{-\alpha-n}\, ds \tag{32}$$

The change of variables $u = (1+s)^{-1}$ in (32) gives (11).

Using the definition of TVaR and Proposition 2.1 with $n = 2$, gives

$$TVaR_2 = \frac{\alpha(1+\alpha)\beta^\alpha}{1-\kappa} \int_{q_2}^\infty \frac{s^2}{(\beta+s)^{\alpha+2}}\, ds$$

from which (13) follows, via the change of variables $u = s/\beta$.

Now let us assume that $\alpha \in \mathbb{N} \setminus \{0, 1\}$. We prove (12) by induction on α.

For $\alpha = 2$, computing (11) provides

$$\mathrm{TVaR}_n = \frac{\beta}{(1-\kappa)\, B(2, n)} B(p_n;\, 1, n+1) = \frac{\beta n(n+1)}{1-\kappa} \times \frac{1-(1-p_n)^{n+1}}{n+1} = \frac{n\beta}{1-\kappa} \left(1-(1-p_n)^{n+1}\right)$$

which corresponds to (12) when replacing α by 2.

Assume now that (12) is satisfied for any $2 \le \alpha \le k$ (and for any $n \ge 1$). Let us prove that it holds for $\alpha = k+1$. It comes back to prove the induction on the following expression of the incomplete β function:

$$B(p_n;\, k, n+1) = \frac{(k-1)!\, n!}{(n+k-1)!} B(p_n;\, 1, n+k) - \sum_{j=1}^{k-1} \frac{(k-1)!\, n!}{(n+k-j)!\, j!}\, p_n^j (1-p_n)^{n+k-j}$$

But, since $B(p_n;\, 1, n+k) = \frac{1}{n+k}\left(1-(1-p_n)^{n+k}\right)$,

$$B(p_n;\, k, n+1) = \frac{(k-1)!\, n!}{(n+k)!} - \sum_{j=0}^{k-1} \frac{(k-1)!\, n!}{(n+k-j)!\, j!}\, p_n^j (1-p_n)^{n+k-j} \tag{33}$$

It is immediate to check that (33) is satisfied for $k = 2$. Then, assuming that (33) holds for $k \in \mathbb{N} \setminus \{0, 1\}$ (and any $n \ge 1$), we check that it remains true for $k+1$. An integration by part gives

$$B(p_n;\, k+1, n+1) = -\frac{1}{n+1} p_n^k (1-p_n)^{n+1} + \frac{k}{n+1} B(p_n;\, k, n+2)$$

Under the inductive assumption, we can apply (33) to express $B(p_n; k, n+2)$; replacing it in the previous equation provides that $B(p_n; k+1, n+1)$ satisfies (33). Hence the result.

- *Proof of Corollary 2.1*
  i. It is immediate knowing the fact that $\mathbb{E}(X) = \frac{\beta}{\alpha-1}$ for X (α, β)-Pareto distributed.
  ii. Combining (10) with (11) in the definition of D_n gives (15).
  For the case $n=2$, (16) can be directly deduced from (15). An alternative, with simpler computation, is to deduce (16) from (13) and the definition (4) of the diversification benefit.

B. Proofs of the results given for the Weibull–Gumbel model

- *Proof of Proposition 2.2*

The proof of this proposition is based on mixing techniques and on the following technical lemma.

Lemma B.1 For $n \geq 1$, we have

$$\partial_s^n\left(e^{-c\sqrt{s}}\right) = \frac{c(-1)^n\Gamma\left(n-\frac{1}{2}\right)}{2\sqrt{\pi}}\, s^{\frac{1}{2}-n}\, e^{-c\sqrt{s}}\, {}_1F_1\left(1-n, 2-2n;\, 2c\sqrt{s}\right)$$

where ∂_s denotes the partial derivative w.r.t. s.

Proof of Lemma B.1. We proceed by induction.

For $n=1$, Lemma B.1 is well satisfied since $\partial_s\left(e^{-c\sqrt{s}}\right) = -\frac{c}{2}s^{-1/2}\, e^{-c\sqrt{s}}$, $\Gamma(1/2) = \sqrt{\pi}$, and ${}_1F_1(0, 0; z) = 1$, for all real z. Suppose now that Lemma B.1 is true for $n \geq 1$ and let us check it for $n+1$. Under the induction assumption, we have

$$\partial_s^{n+1}\left(e^{-c\sqrt{s}}\right) = \frac{c\,(-1)^n\Gamma\left(n-\frac{1}{2}\right)}{2\sqrt{\pi}}\, \partial_s\left(s^{\frac{1}{2}-n}\, e^{-c\sqrt{s}}\, {}_1F_1\left(1-n, 2-2n;\, 2c\sqrt{s}\right)\right)$$

Let us compute $A_n(s) := \partial_s\left(s^{\frac{1}{2}-n}\, e^{-c\sqrt{s}}\, {}_1F_1(1-n, 2-2n;\, 2c\sqrt{s})\right)$, using the following properties of hypergeometric functions (see, e.g. Gradshteyn, 1988):

i. $\partial_z({}_1F_1(a, b;\, z)) = \frac{a}{b}\, {}_1F_1(a+1, b+1;\, z)$
ii. ${}_1F_1(a, 2a;\, z) = e^{z/2}\, {}_0F_1\left(a+\frac{1}{2};\, \frac{z^2}{16}\right)$
iii. ${}_1F_1(a, 2a-1;\, z) = \frac{1}{4}\, e^{z/2}\left(4\, {}_0F_1\left(a-\frac{1}{2};\, \frac{z^2}{16}\right) + z\,\frac{\Gamma\left(a-\frac{1}{2}\right)}{\Gamma\left(a+\frac{1}{2}\right)}\, {}_0F_1\left(a+\frac{1}{2}, \frac{z^2}{16}\right)\right)$
iv. ${}_0F_1(a;\, z) = e^{-2\sqrt{z}}\, {}_1F_1\left(a-\frac{1}{2}, 2a-1;\, 4\sqrt{z}\right)$
where ${}_0F_1(a; x)$ is the confluent hypergeometric function defined by ${}_0F_1(a;\, z) = \sum_{k=0}^{\infty}\frac{1}{(a)_k}\frac{z^k}{k!}$ with $(a)_k = \frac{\Gamma(a+k)}{\Gamma(a)}$

We have

$$A_n(s) = \frac{1}{2}\, s^{-n}e^{-c\sqrt{s}}\left[c\, {}_1F_1\left(2-n,\, 3-2n;\, 2c\sqrt{s}\right) - \left(c+(2n-1)s^{-1/2}\right){}_1F_1\left(1-n,\, 2-2n;\, 2c\sqrt{s}\right)\right]$$

$$= \frac{1}{2}\,(1-2n)\, s^{-n-1/2}\, {}_0F_1\left(\frac{1}{2}-n;\, \frac{c^2 s}{4}\right) = \frac{1}{2}(1-2n)\, s^{-n-1/2}\, e^{-c\sqrt{s}}\, {}_1F_1\left(-n,\, -2n;\, 2c\sqrt{s}\right)$$

using (i) in the first equality, then both (ii) and (iii) in the second one, and (iv) in the last one. We can conclude that Lemma B.1 is satisfied for $n+1$. □

To prove Proposition 2.2, we will also need the following relation.

For $n \geq 0$ and $a, b \in \mathbb{R}$ (see, e.g. Albano *et al.*, 2011), we have

$$I_{n,b}(a) = \int_0^\infty x^{n-1/2} e^{-ax-b/x} dx = (-1)^n \partial_a^n \big(I_{0,b}(a)\big)$$

$$\text{with} \quad I_{0,b}(a) = \int_0^\infty x^{-1/2} e^{-ax-b/x}\, dx = \sqrt{\frac{\pi}{a}} e^{-2\sqrt{ab}} \tag{34}$$

when differentiating $I_{n,b}(a)$ with respect to the parameter a.

Let us prove now Proposition 2.2. As given in Marshall & Olkin (1988), it is enough to choose $H(x) = e^{-x}$ $(x > 0)$ and Θ a Lévy-distributed $(0, c^2/2)$ positive rv with $c > 0$, and pdf $f_\Theta(x) = \frac{c}{2\sqrt{\pi x^3}}\, e^{-\frac{c^2}{4x}}$, for $x \geq 0$, then to apply the Oakes–Marshall–Olkin theorem.

Computing the Laplace transform of Θ via the change of variables $u = \theta^{-1}$, gives

$$L_\Theta(t) = \int_0^\infty e^{-\theta t} f_\Theta(\theta)\, d\theta = I_{0,t}\left(c^2/4\right) = e^{-c\sqrt{t}}$$

So, with this choice of H, the marginal distribution defined in (3) satisfies $L_\Theta(-\ln H(x)) = e^{-c\sqrt{x}}$. It corresponds on $(0, \infty)$ to the survival cdf of a Weibull $(c, 1/2)$ defined in (17). Hence we obtain the marginal distributions of our model.

Now, since the generator of the Archimedean survival copula defined in (2), is given by

$$\phi(t) = L_\Theta(t)^{-1} = \frac{1}{c^2}(-\ln(t))^2$$

we deduce the structure of dependence assumed in our model, namely (19) when taking the parameter $\theta = 2$.

For the 2nd part of the proof, we can write, as in the proof of Proposition 2.1,

$$f_n(s) = \int_0^\infty f_{S_n|\Theta}(s)\, f_\Theta(\theta) d\theta = \frac{c\, s^{n-1}}{2\sqrt{\pi}\,\Gamma(n)} \int_0^\infty \theta^{(n-1)-1/2}\, e^{-s\theta-\frac{c^2}{4\theta}} d\theta \tag{35}$$

Therefore, using (34) provides

$$f_n(s) = \frac{c\, s^{n-1}}{2\sqrt{\pi}\,\Gamma(n)} (-1)^{n-1} \sqrt{\pi}\, \partial_s^{n-1}\left(s^{-1/2} e^{-c\sqrt{s}}\right) = \frac{s^{n-1}}{\Gamma(n)} (-1)^n \partial_s^n\left(e^{-c\sqrt{s}}\right)$$

Applying Lemma B.1 allows to conclude.

- *Proof of Theorem 2.2*
  By definition of the TVaR, we can write, using the expression (35) for f_n

$$\text{TVaR}_n = \frac{1}{1-\kappa} \int_{q_n}^\infty s\, f_n(s)\, ds = \frac{c}{2\sqrt{\pi}\,(1-\kappa)\,\Gamma(n)} \int_0^\infty \theta^{n-3/2}\, e^{-\frac{c^2}{4\theta}} \int_{q_n}^\infty s^n\, e^{-s\theta}\, ds\, d\theta$$

  But we have, by the change of variable $t = s\theta$

$$\int_{q_n}^\infty s^n\, e^{-s\theta}\, ds = \frac{1}{\theta^{n+1}}\, \Gamma(n+1\,;\ \theta\, q_n) = \frac{n!\, e^{-\theta q_n}}{\theta^{n+1}} \sum_{k=0}^{n} \frac{(\theta\, q_n)^k}{k!}$$

$\Gamma(s;\, x)$ denoting the incomplete γ function defined by $\int_x^\infty t^{s-1}e^{-t}dt$ that can be expressed as a discrete sum (see Gradshteyn, 1988).
So we deduce that

$$\mathrm{TVaR}_n = \frac{c\,n}{2(1-\kappa)\sqrt{\pi}}\left(\int_0^\infty (1+\theta q_n)\,\theta^{-5/2}\,e^{-\frac{c^2}{4\theta}-\theta q_n}d\theta + \sum_{k=2}^{n}\frac{q_n^k}{k!}I_{k-2,c^2/4}(q_n)\right)$$
$$= \frac{c\,n}{2(1-\kappa)\sqrt{\pi}}\left(I_{1,q_n}(c^2/4) + q_n I_{0,q_n}(c^2/4) + \sum_{k=2}^{n}\frac{q_n^k}{k!}I_{k-2,c^2/4}(q_n)\right)$$

using the definition (34) in the first equation, and the change of variable $\theta = u^{-1}$ to compute the integral, in the second one.

On one hand, we have $I_{0,q_n}(c^2/4) = \frac{2\sqrt{\pi}}{c}\,e^{-c\sqrt{q_n}}$; on the other hand, a straightforward computation gives $I_{1,q_n}(c^2/4) = \frac{4\sqrt{\pi}}{c^3}\,e^{-c\sqrt{q_n}}\left(1+c\sqrt{q_n}\right)$. Moreover, we can write, using once again (34)

$$I_{k-2,c^2/4}(q_n) = (-1)^k\partial_{q_n}^{k-2}\left(I_{0,c^2/4}(q_n)\right) = (-1)^{k-1}\frac{2\sqrt{\pi}}{c}\,\partial_{q_n}^{k-1}\left(e^{-c\sqrt{q_n}}\right)$$

from which we deduce, applying Lemma B.1, that

$$\sum_{k=2}^{n}\frac{q_n^k}{k!}\,I_{k-2,c^2/4}(q_n) = q_n^{3/2}\,e^{-c\sqrt{q_n}}\sum_{k=2}^{n}\frac{\Gamma\left(k-\frac{3}{2}\right)}{k!}\,{}_1F_1\left(2-k,\,4-2k;\,2c\sqrt{q_n}\right)$$

Combining these results provide (22).

Applying (22) with $n=1$ provides $TVaR_1 = \frac{2e^{-c\sqrt{q_1}}}{(1-\kappa)c^2}\left(1+c\,q_1^{1/2}+\frac{c^2q_1}{2}\right)$, from which (23) follows since $q_1 = (\ln(1-\kappa))^2/c^2$. For $n=2$, (24) is a direct application of (22).

- *Proof of Corollary 2.2*

  i. It is immediate knowing that $\mathbb{E}(X) = c\Gamma\left(1+\frac{1}{\tau}\right) = 2c$ when X is $(c,\,\tau)$-Weibull distributed.

  ii. Combining (22), (24), respectively, with (23) in the definition of the diversification benefit (4) gives the results.